1971

A comparison of ratios and deviations for expressing weaning weights and grades in beef cattle records

Jack Atha Richey
Iowa State University

71-21,970

RICHEY, Jack Atha, 1935-
A COMPARISON OF RATIOS AND DEVIATIONS FOR EXPRESSING WEANING WEIGHTS AND GRADES IN BEEF CATTLE RECORDS.

Iowa State University, Ph.D., 1971
Agriculture, animal culture

A comparison of ratios and deviations

for expressing weaning weights and grades

in beef cattle records

by

Jack Atha Richey

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1971

TABLE OF CONTENTS

INTRODUCTION

Most beef cattle performance programs use what is termed a "weight ratio" or a ratio of an observation to a group average to compare calves in different groups and to compare dams on their progeny which come from different groups. Usually the group is contemporary in time, management and sex. Genetic theory is developed for the comparison of individuals using deviations from such group averages. Statistical properties of these deviated records are reasonably well established, while the statistical properties of ratios have received little attention. Therefore, the statistical properties of ratios need study so that the two systems can be compared and the most appropriate method for expressing beef records chosen.

The purposes of this study are as follows:

1. To investigate the statistical properties of the ratio.
2. To examine the consequences on the ratio of adjusting group averages for numbers.
3. To determine the appropriateness of the ratio in comparison to the deviation as a means of adjusting weaning weight and grade of beef cattle to maximize the relative importance of genetic differences.

4. To ascertain the consequences of using ratios to compare the genetic merit of individuals.
   a) To study the use of ratios for sire evaluation in the context of the analysis of variance.
   b) To develop appropriate prediction formula for estimating most probable producing ability using records expressed as ratios.

LITERATURE REVIEW

The use of ratios of observations to express beef cattle records apparently originated with Dr. T. C. Cartwright and Dr. Bruce Warwick of Texas A. and M. University. They first expressed gain as a percent of group average when reporting performance at the McGregor, Texas station in 1953. From this beginning, the ratio concept was developed into gain ratio and has been extended to weaning weights, yearling weights and other traits of beef cattle (Maddox, 1970, private communication). The original purpose of expressing records as ratios was to give a means of comparison across sexes and years. The ratio was never intended to be used in statistical analyses, but was a method of reporting research work to the public (Cartwright, 1970, private communication).

There appears to be no published data in which this particular type of ratio was investigated or compared to the deviation.

Two areas of concern in the study of the ratio have received attention. They are the use of multiplicative correction factors since the ratio is a form of multiplicative adjustment and the use of deviated records in the context of the analysis of variance for estimation of variance components. These areas will be reviewed.

The most common criteria used in determining the appropriateness of adjustment factors have been the equalization of means and within group variances (Koch, et al., 1959; Brinks, et al., 1961; Cundiff, 1966). Adjustments for differences in group means can be made by adding a constant, but this type of adjustment does not change the variance within groups. A multiplicative adjustment is appropriate if the coefficients of variation among groups are equal. In this case a ratio of group means is the desired adjustment, since both the means and variances are equalized. In studies where additive and multiplicative correction factors were actually compared, multiplicative sex correction factors better equalized weaning weight variances within sexes than did additive correction factors (Brinks, et al., 1961; Cundiff, 1966; Sellers, 1968). However, Cundiff (1966) found that the variances within age of dam classes became more unequal when multiplicative factors were applied. Cundiff (1966) but not Sellers (1968) found that a multiplicative sex adjustment also adjusted for the sex by management interaction.

Many investigators have expressed dairy records as deviations from various contemporary averages because of the computational difficulties of analyzing large numbers of records according to complex models (Van Vleck, et al., 1961). Therefore, expected values of sums of squares and cross products of deviations were calculated for a one-way classification analysis in terms of a more complex model. These expectations were then compared with analyses of actual data expressed as deviations. Ratios of total variances from different models were calculated from the data and compared to theoretical expected ratios from the same models. The ratios of variance from the data and the theoretical models were close enough to conclude that the models and expected values effectively described the data, and that deviated records could be utilized in analyzing dairy records. It was, however, noted that any investigator who plans to analyze deviations should understand the expectations and assumptions involved.

The analysis of variance was introduced by Sir Ronald A. Fisher and is essentially an arithmetic process for partitioning a total sum of squares into components associated with recognized sources of variations (Steel and Torrie, 1960). A basic understanding of the assumptions and the consequences of failure to meet the assumptions for analysis of variance are necessary in determining whether ratios may reasonably be used when analyzing beef cattle data. Therefore, the assumptions underlying analysis of variance procedures for estimating variance components, some consequences when the assumptions are not satisfied, and the use of transformations to enable data to more nearly meet the assumptions for the analysis of variance will be reviewed.

The analysis of variance procedures ordinarily used to estimate components of variance, to test hypothesis and to infer properties of the population from which the data were drawn, require the fulfillment of certain assumptions if the inferences are to be valid (Eisenhart, 1947). The necessary assumptions for component of variance estimation are: 1) The observed values are random variables that are distributed about a common mean which is a fixed constant. 2) The random variables are sums of component random variables. These two assumptions also imply that the mean values of the random variables are zero. 3) The random variables are distributed with homogenous variances and all covariances among them are zero. 4) The component random variables are all normally distributed. This assumption is necessary if exact tests of significance are needed. When these assumptions are all satisfied, then all of the analysis of variance procedures for estimating and testing to determine whether to infer the existence of components of variance are strictly valid.

Some consequences when the assumptions for the analysis of variance are not satisfied were discussed by Cochran (1947). Listed among the factors most likely to cause severe disturbances in analysis of variance procedures and give misleading results or produce a serious loss of information were the presence of gross errors, marked departure from the additive relationship, and changes in the error variance, either related to the mean or to certain treatments or parts of the experiment. The principle method suggested to improve analyses were the omission of certain observations, treatments, or replications, subdivision of the error variance, and transformation to another scale before analysis.

Bartlett (1947) discussed the use of transformations with particular reference to the analysis of variance. He stated that the usual purpose of the transformation is to change the scale of the measurements in order to make the analysis more valid. The conditions required for assessing accuracy in the ordinary analysis of variance include the important one of a constant residual or error variance, and if the variance tends to change with the mean level of the measurements, the variance will only be stabilized by a suitable change of scale. Reference to the ideal case suggested: a) The variance of the transformed variate should be unaffected by changes in the mean level. b) The transformed variate should be normally distributed. c) The transformed scale should be one for which an arithmetic average is an efficient estimate of the true mean level for any particular group of measurements. d) The transformed scale should be one for which real effects are linear and additive.

When we estimate the ratio to a group average or the deviation from a group average, we assume we have measured the group mean without error. But in fact, the confidence we put in the group average is dependent upon the number of observations from which the mean is calculated. A method for adjusting group averages for small numbers of observations was demonstrated by Heidhues, et al. (1961).

In this study, the trait in question is considered to be a continuous variable normally distributed with mean (μ) and variance (σ^2). The group averages are means of normally distributed variables and are themselves normally distributed, but have different means and variances. A joint function of the variables and their group averages are distributed according to the bivariate normal. Given this situation, Heidhues determined a "best" estimate of the true group average given an estimated or observed group average. The group or stablemate average was of the following form:

The model can be expressed as:

$$y_{ij} = \mu + h_i + w_{ij}$$

where

μ = age corrected five year breed average plus fixed effects in the group,

h_i = a random herd effect, and

w_{ij} = a random environmental effect plus a genetic effect.

The group average can be estimated as:

$$\sum_{j=1}^{n} \frac{y_{ij}}{n} = \bar{y} \text{ and the } E(\bar{y}) = \mu$$

Henderson (1948) provided the basis for the adjustment by showing that an estimate of the true group average could be obtained by adjusting the group average for random effects. In order to obtain a better estimate of the true average of a group, Henderson indicated the need to find the mean of the conditional density function

$$f(\bar{x}/\bar{y}) = \frac{1}{2\,\pi\,\sigma_{\bar{x}}^2\,(1-r^2)}\, e^{-\left[\frac{1}{2(1-r^2)\sigma_{\bar{x}}^2}\,[\bar{x} - \mu_{\bar{x}} - r\,\frac{\sigma_{\bar{x}}}{\sigma_{\bar{y}}}(\bar{y} - \mu_{\bar{y}})\,]^2\right]}$$

where

r = correlation between $\bar{x}$ and $\bar{y}$,

$\sigma_{\bar{x}}^2$ = variance of the true group average, and

$\sigma_{\bar{y}}^2$ = variance of the estimated group average.

The mean of the conditional distribution is

$$\text{mean} = \mu_{\bar{x}} + r\,\frac{\sigma_{\bar{x}}}{\sigma_{\bar{y}}}\,(\bar{y} - \mu_{\bar{y}}) = \bar{y}_a$$

where

$\bar{x}$ = the true group average,

$\bar{y}$ = the estimate of the group average, and

$\bar{y}_a$ = the adjusted or best estimate of the true group average.

The adjusted average may also be expressed as

$$\bar{y}_a = \mu_{\bar{x}} + \frac{\sigma_{\bar{x}\bar{y}}}{\sigma_{\bar{x}}\sigma_{\bar{y}}}\,\frac{\sigma_{\bar{x}}}{\sigma_{\bar{y}}}\,(\bar{y} - \mu_{\bar{y}})$$

$$= \mu_{\bar{x}} + \frac{\sigma_{\bar{x}\bar{y}}}{\sigma^2_{\bar{y}}} (\bar{y} - \mu_{\bar{y}})$$

$$= \mu_{\bar{x}} + b(\bar{y} - \mu_{\bar{y}}) \quad \text{where} \quad b = \frac{\sigma_{\bar{x}\bar{y}}}{\sigma^2_{\bar{y}}} .$$

The covariance $(\bar{x}, \bar{y}) = E[(\mu + h_i)(\mu + h_i + \sum_j w_{ij})]$

$= \sigma^2_h$, and the variance of $\bar{y}$ is

$$\mathrm{Var}(\mu + h_i + \sum_j \frac{w_{ij}}{n}) = \sigma^2_h + \frac{\sigma^2_e}{n} .$$

$$\text{Therefore, } b = \frac{\sigma^2_h}{\sigma^2_h + \frac{\sigma^2_w}{n}} = \frac{n}{n + \frac{\sigma^2_w}{\sigma^2_h}}$$

$$\text{and } \bar{y}_a = \mu_{\bar{x}} + \frac{n}{n + \frac{\sigma^2_w}{\sigma^2_h}} (\bar{y} - \mu_{\bar{y}}) .$$

If we assume that $\mu_{\bar{y}}$ is measured without error then $\mu_{\bar{y}} = \mu_{\bar{x}}$ and

$$\bar{y}_a = \mu_{\bar{x}} + \frac{n}{n + \frac{\sigma^2_w}{\sigma^2_h}} (\bar{y} - \mu_{\bar{x}}) .$$

In summary, ratios have been used extensively in reporting beef cattle performance and in performance programs with no reported investigations of the validity of this type of adjustment. Multiplicative correction factors have been shown to be more appropriate than additive corrections for sex since they better equalized mean and variances within classes, while for other factors additive corrections appear more appropriate. In the special case where the coefficients of variation are equal, a ratio of the group means is the appropriate adjustment.

Deviations have been examined more thoroughly than ratios. The statistical properties of deviations and the advantages and disadvantages of using deviations for the adjustment of data have been investigated.

The appropriateness of ratios or deviations relative to the assumptions underlying analysis of variance procedures need investigation.

The confidence placed in estimates of group means is dependent on the number of observations in the group. A method of adjusting group averages for small numbers of observations is available.

DATA

The data were weaning weight and weaning grade on 28,545 calves from 203 herds in Iowa representing six breeds. The data were collected by the Iowa Beef Improvement Association (I.B.I.A.) over a thirteen year period, 1956-1968, inclusive. Only those calves weaned between 160 and 250 days of age were used in this study. The data were divided into breed, herd, year, sex, season and management groups. There were 2,478 such groups having two or more observations. Groups with only one record were removed. Actual birth weights were used when available or were assigned when actual weights were not taken. The assigned weights were 60 pounds for Angus, 70 pounds for Hereford and Shorthorn and 80 to 85 pounds for Charolais calves. Weaning weights were adjusted to a 205-day standard by multiplying the average daily gain of a calf from birth to weaning by 205 and adding the birth weight. Age of dam adjustments for weaning weight were computed by multiplying the computed 205-day record by the adjustment factors recommended by the United States Beef Cattle Records Committee Report (1965).

Weaning grades were a visual appraisal based on a 17-point scale where each point represented one-third of a feeder grade. A value of 13 represented an "average choice" quality score based on USDA feeder calf grades.

Management referred to the presence or absence of creep feed. To be classified as creep-fed, calves must have had access to creep feed for at least six weeks.

The six breeds were Angus, Hereford, Shorthorn, Charolais, Red Angus and Polled Hereford. The four seasons were winter (Dec.-Feb.), spring (March-May), summer (June-Aug.) and fall (Sept.-Nov.). A record was classified for season by month of birth, not month of weaning. Sexes were bulls, steers and heifers.

ANALYSIS OF DATA

A preliminary analysis was conducted on 12 years' data which included 19,905 observations divided into 1,878 groups. The initial items of information needed were group statistics. Number of observations, average, variance and coefficient of variation for weaning weight and weaning grade were calculated for each group. These group statistics were analyzed by least squares procedures to examine the influence of sex, season, management and the two-way interactions. Since inferences were to be made across herds and years which were considered to be random, only the fixed effects of sex, season, management and their interactions were included in the model for this analysis. The model was

$$y_{ijkl} = \mu + s_i + m_j + x_k + mx_{jk} + sx_{ik} + sm_{ij} + e_{ijkl},$$

where

$i = 1, 4;\ j = 1, 2;\ k = 1, 3;$

y_{ijkl} = group number, mean, variance or coefficient of variation for weaning weight or weaning grade for the l^{th} group of the k^{th} sex, j^{th} management and the i^{th} season;

μ = the overall mean for the dependent variable,

s_i = deviation from μ due to the i^{th} season,

m_j = deviation from μ due to the j^{th} management,

x_k = deviation from μ due to the k^{th} sex,

mx_{jk} = deviation from μ due to the interaction of the j^{th} management and the k^{th} sex,

sx_{ik} = deviation from μ due to the interaction of the i^{th} season and the k^{th} sex,

sm_{ij} = deviation from μ due to the interaction of the i^{th} season and the j^{th} management, and

e_{ijkl} = random error.

This procedure gave least squares means for the group statistics by sex, season, management and their interactions. Such an analysis of variances and coefficients of variation may not be strictly valid, but they do allow comparisons to be made on the relative equality of the statistics over the factors in the model.

The data were then analyzed within breed, sex, season, and management factors for an estimate of the within and among herd variance components. The ratio of within herd variance to among herd variance was used in the calculation of the adjustment of group averages for small numbers according to the method of Heidhues, et al. (1961). The adjustment was

$$\bar{y}_a = \mu_{\bar{x}} + \frac{n_i}{n_i + \frac{\sigma_w^2}{\sigma_h^2}} (\bar{y} - \mu_{\bar{x}})$$

where

$\bar{y}_a$ = the adjusted group average of weight or grade,

$\mu_{\bar{x}}$ = 5-year breed, season, sex and management group average,

n_i = the number of observations in a group,

σ_w^2 = the within herd variance,

σ_h^2 = the among herd variance, and

$\bar{y}$ = the group average for weight or grade.

The group averages were regressed on five year averages for breed, season, sex and management. The adjustment was accomplished by reading in the data from tape and computing group averages which were stored on a disk while the additional data accumulated for the breed, season, sex and management averages. These averages could be stored internally because of the smaller number involved. When all of the data for a breed had been read and the averages computed and stored in an array internally and all of the group averages computed and stored on a disk, then the disk was rewound and read. As each group average was read from the disk the appropriate average was called from the array and the group average was regressed. The data for the groups were accumulated by years such that a particular group was regressed on an average containing information from that year and the four previous years if available. The data from the first year were regressed on breed averages for that year only and each succeeding year on cumulative averages until the sixth year when the data from the first year were dropped from the averages. Only the fifth and succeeding years were regressed on five year averages.

The data were expressed as observations, ratios of observations to the group average, ratios of the observations to the group average adjusted for number of observations in the group, deviations of observations from the group average and deviations of observations from the adjusted group average.

The model used to describe an observation in a group was

$$y_{ijklmno} = \mu + b_i + h_{ij} + a_k + s_l + x_m + m_n + \text{interactions} + e_{ijklmno},$$

where

$y_{ijklmno}$ = weaning weight or weaning grade for the o^{th} calf in the n^{th} management group of the m^{th} sex, born in the l^{th} season of k^{th} year in the j^{th} herd of the i^{th} breed,

μ = mean weaning weight or grade,

b_i = deviation from μ due to the i^{th} breed,

h_{ij} = deviation from μ due to the j^{th} herd in the i^{th} breed,

a_k = deviation from μ due to the k^{th} year,

s_l = deviation from μ due to the l^{th} season,

x_m = deviation from μ due to the m^{th} sex,

m_n = deviation from μ due to the n^{th} management,

interactions = deviation from μ of all the possible interaction effects, and

$e_{ijklmno}$ = random error.

Differences among group means were examined in an hierarchal analysis of variance. For this analysis the model was completely nested as management within sex within season within year within herd within breed. Since main effects were not being estimated, this was a convenient method for determining whether any component for a main effect plus all interactions with higher order elements in the model accounted for a significant percentage of the variance. If expressing the data as ratios and deviations removed group effects, then the only significant source of variation should be due to $e_{ijklmno}$. If significant group effects were present, the main effects and two-way interactions could be examined in the cross classification model. The hierarchal analysis of variance was as shown in Table 1.

Table 1. Sources of variation in the hierarchal analysis of variance

Source	df
Breed	5
Herd/b	189
Year/h/b	455
Season/y/h/b	601
Sex/s/y/h/b	1110
Management/Sex/s/y/h/b	117
Within management	24278

A cross classification model including the fixed effects of breed, sex, season and management plus two-way interactions among sex, season and management and the interactions of breed with sex and season was used to analyze the equality among the group variances for the five different methods of expressing the data. These group variances were analyzed by least squares. The least squares analyses were as shown in Table 2.

Table 2. Sources of variation in the least squares analyses

Source	df
Breed	5
Season	3
Sex	2
Management	1
Breed x sex	9
Breed x management	4
Season x sex	6
Season x management	3
Sex x management	2
Remainder	2442

The failure of the degrees of freedom for breed by sex and breed by management to be as large as expected was due to small numbers of Red Angus calves and their failure to be represented in all sex and management groups.

The decision of whether to use ratios or deviations to express weaning weights and grades when comparing individuals from different groups was based on which method best removed effects due to the group means, and at the same time equalized variances across groups. The analyses to this point were designed to answer this question.

After observing results from the previous analyses, attention was given to checking the validity of using ratios in existing formulae for genetic analyses and where these were inappropriate, developing new formulae.

RESULTS

Statistical properties of ratios

A primary assumption used in these analyses was that the group averages (the denominators of the ratios) were measured without error and could, in fact, be treated as constants in the computations. Without this assumption the ratios are nonlinear forms for which expected values are defined but for which no exact methods for estimating variances are available. The condition which lets the denominator be treated as a constant is that the sum of the random effects be zero when summed over all observations in a group. This assumption that the random effects are deviations about a mean of zero is consistent with standard analysis of variance procedures and leads to expected values which are identical with the numerical solutions obtained when computing sums of squares, where the variable is expressed as a ratio to its group average.

It is not implied that these ideas and assumptions are applicable to other types of ratios, but that they do hold for the special ratio of an observation to the group average. Consideration was given to the fact that ratios of random variables may follow the Cauchy distribution which has neither mean nor variance, and that approximate methods for estimating the variance of other ratios are available. The approximate method for estimating the variance of a ratio based on partial derivatives of a Taylor series expansion requires that the numerator and the denominator be identically and independently distributed in order to obtain estimates of variance numerically equal to the square of the coefficient of variation which is obtained through analysis of variance procedures. In addition this method is only appropriate when the data are balanced (Kempthorne, 1969).

Since it is apparent that ratios of observations to the group average have an arithmetic mean of one and since the data were not balanced or independently distributed, neither the Cauchy distribution nor the approximate methods for calculating variances of ratios were particularly helpful when examining this particular ratio.

The basic ideas used to explore these data can be easily explained using the following simple models. These models also provided the basis for the solutions to the expected mean squares shown later.

In the simplest form the model can be expressed as

$$x_{ij} = u + g_i + w_{ij}$$

where

x_{ij} = the observed value for a trait measured on the j^{th} individual in the i^{th} group,

u = the mean value for the trait,

g_i = the fixed effects for the i^{th} group, and

w_{ij} = the random effects for the j^{th} individual in the i^{th} group.

The expected value of the fixed effects are $E(g_i) = g_i$ and the expected values of the random effects are $E(w_{ij}) = 0$. The group average can be shown as

$$\bar{x}_{i.} = \sum_j^n \frac{x_{ij}}{n} = \frac{1}{n} \sum_j^n (\mu + g_i + w_{ij}) = \mu + g_i + \sum_j^n \frac{w_{ij}}{n} .$$

The ratio is $r_{ij} = \dfrac{x_{ij}}{x_{i.}} = \dfrac{\mu + g_i + w_{ij}}{\mu + g_i + \sum_j \frac{w_{ij}}{n}}$, a non-linear form. But if $\sum_j^n \frac{w_{ij}}{n} = 0$, then $r_{ij} = \dfrac{\mu + g_i + w_{ij}}{\mu + g_i}$, a weighted linear form for which expected values are defined, and which has a predictable form of variance.

As a comparison the deviation can be expressed as

$$d_{ij} = (x_{ij} - \bar{x}_{i.}) = (\mu + g_i + w_{ij} - \mu - g_i - \sum_j \frac{w_{ij}}{n})$$

$$= w_{ij} - \sum_j^n \frac{w_{ij}}{n} \text{ or if } \sum_j^n \frac{w_{ij}}{n} = 0$$

$$= w_{ij} .$$

The expected values are

$$E(r_{ij}) = E(\frac{\mu + g_i + w_{ij}}{\mu + g_i}) = E(1 + \frac{w_{ij}}{\mu + g_i}) = 1$$

and

$$E(d_{ij}) = E(w_{ij}) = 0 \ .$$

The variances are

$$\mathrm{Var}(r_{ij}) = E[r_{ij} - E(r_{ij})]^2 = E\left[1 + \frac{w_{ij}}{\mu + g_i} - 1\right]^2$$

$$= \frac{\sigma_w^2}{(\mu + g_i)^2} = \frac{\sigma_w^2}{\bar{x}_{i.}^2} = \text{the coefficient of}$$

variation squared. And,

$$\mathrm{Var}(d_{ij}) = E[w_{ij}]^2 = \sigma_w^2 \ .$$

Therefore, for the particular ratio of an observation to a group average the variance of the ratios in a particular group can be estimated as the square of the coefficient of variation for that group. Standard analysis of variance procedures to estimate the variance of ratios yield an algebraic identity of the square of the coefficient of variation when computed for each group. The algebra is as follows:

$$\mathrm{Var}(r_{ij}) = \mathrm{Var}\left[\frac{x_{ij}}{\bar{x}_{i.}}\right] = \frac{\dfrac{\sum\limits_j (x_{ij})^2}{\bar{x}_{i.}^2} - \dfrac{\dfrac{(\sum\limits_j x_{ij})^2}{\bar{x}_{i.}^2}}{N}}{N - 1}$$

$$\text{Coeff. of Var}(x_{ij})^2 = \frac{s_i^2}{\bar{x}_{i.}^2} = \frac{\dfrac{\sum_j x_{ij}^2 - \dfrac{(\sum_j x_{ij})^2}{N}}{N-1}}{\bar{x}_{i.}^2}$$

$$= \frac{\dfrac{N \sum_j x_{ij}^2 - (\sum_j x_{ij})^2}{N \bar{x}_{i.}^2}}{N-1}$$

$$= \frac{\dfrac{\sum_j x_{ij}^2}{\bar{x}_{i.}^2} - \dfrac{(\sum_j x_{ij})^2}{N\bar{x}_{i.}^2}}{N-1}$$

$$= \frac{\dfrac{\sum_j (x_{ij})^2}{\bar{x}_{i.}^2} - \dfrac{\dfrac{(\sum_j x_{ij})^2}{\bar{x}_{i.}^2}}{N}}{N-1} = \text{Var}\ \frac{x_{ij}}{\bar{x}_{i.}}$$

Preliminary analyses

The preliminary least squares analyses on 19,905 calves in 1,978 groups included those group variables found in Table 3. Table 3 also gives the estimated means and standard deviations for the variables.

Table 3. Overall means and standard deviations for dependent variables in the preliminary least squares analyses

Variable	Mean	SD
Number/group	10.390	58.045
Average weight/group	419.077	58.045
Variance of weight/group	2264.659	2529.763
Coefficient of variation for weight/group	10.294	5.325
Average grade/group	13.821	2.175
Variance of grade/group	1.364	2.860
Coefficient of variation for grade/group	7.033	5.660

The least square means, variances and coefficients of variation for weaning weight and weaning grade are shown for each class of the independent variables considered in the preliminary analyses of variance in Table 4. The effects of sex, management and the two-way interactions among these elements of the model on the group statistics of numbers, means, variances and coefficients of variation are summarized in Table 5. Numbers were significantly different ($P < .05$) in management groups and in season and sex by management groups ($P < .01$). Average number per group was generally higher for noncreep fed calves and for spring born calves, although creep-fed bulls and noncreep-fed heifers comprised the largest sex by management groups. Weaning weight averages were significantly different ($P < .005$) for sex, season, management and sex by management groups, and

Table 4. Least squares means, variances and coefficients of variation for weaning weight and weaning grade

Source of variation	Number of groups	Mean	Variance	Coefficient of variation
Weaning Weight				
Sex				
Bulls	707	439.63	2456.51	10.17
Steers	264	423.00	2357.31	9.91
Heifers	901	394.85	2006.14	10.17
Management				
Creep	999	444.56	2465.18	9.74
Noncreep	873	393.76	2081.46	10.43
Season				
Winter	171	430.32	1683.46	8.24
Spring	1040	425.06	2245.47	10.37
Summer	458	407.24	2685.04	10.87
Fall	203	414.02	2479.29	10.85
Weaning Grade				
Sex				
Bulls	707	13.88	1.51	7.43
Steers	264	13.44	1.44	7.41
Heifers	901	13.98	1.24	6.75
Management				
Creep-fed	999	14.19	1.35	6.74
Noncreep-fed	873	13.34	1.45	7.65
Season				
Winter	171	14.00	1.26	7.22
Spring	1040	13.69	1.31	7.08
Summer	458	13.56	1.64	7.04
Fall	203	13.82	1.38	7.44

Table 5. Preliminary least squares analyses

Source of variation	df	MS	F-value
Number per group			
Sex	2	372.955	1.27
Management	1	1765.893	6.01*
Season	3	8630.279	29.39***
Sex x management	2	2212.363	7.54**
Sex x season	6	305.624	1.04
Management x season	3	646.962	2.20
Remainder	1854	293.619	
Average weaning weight			
Sex	2	247444.264	111.70***
Management	1	541932.623	249.16***
Season	3	30289.819	13.67***
Sex x management	2	38123.621	17.21***
Sex x season	6	970.670	0.44
Management x season	3	9554.578	4.31*
Remainder	1854	2215.201	
Weaning weight variance			
Sex	2	25814200.203	4.15*
Management	1	31480689.151	5.06*
Season	3	27256475.917	4.38**
Sex x management	2	14408212.101	2.31*
Sex x season	6	16912776.413	2.72*
Management x season	3	818158.802	0.13
Remainder	1854	6228064.677	
Coefficient of variation for weight			
Sex	2	2.907	0.10
Management	1	102.853	3.68*
Season	3	143.032	5.12**

*P < .05.

**P < .01.

***P < .005.

Table 5. (Continued)

Source of variation	df	MS	F-value
Coefficient of variation for weight (cont)			
Sex x management	2	11.270	0.40
Sex x season	6	68.645	2.46*
Management x season	3	3.105	0.11
Remainder	1854	27.957	
Average weaning grade			
Sex	2	24.844	2.76
Management	1	150.989	33.58***
Season	3	14.806	1.10
Sex x management	2	3.220	0.36
Sex x season	6	2.310	0.09
Management x season	3	3.360	0.25
Remainder	1854	8335.793	
Weaning grade variance			
Sex	2	8.512	1.04
Management	1	2.181	0.27
Season	3	9.867	1.21
Sex x management	2	30.250	3.71*
Sex x season	6	5.928	0.73
Management x season	3	6.416	0.79
Remainder	1854	8.159	
Coefficient of variation for grade			
Sex	2	62.333	1.96
Management	1	176.623	5.56**
Season	3	6.099	0.19
Sex x management	2	81.382	2.56
Sex x season	6	34.043	1.07
Management x season	3	52.694	1.66
Remainder	1854	31.758	

significantly different ($P < .05$) for management by season groups. Variances were proportional to the means for sex and management groups, but were inversely proportional to the means for season groups. Coefficients of variation were more nearly equal than were variances for weaning weight.

Means for weaning grade were significantly different only for management groups ($P < .005$). Variances for weaning grades were significantly different only for sex by management groups ($P < .05$), and coefficients of variation were significantly different only for management ($P < .01$).

Adjustment for number of observations per group

The hierarchal analyses of variance used to establish the ratio of within herd variance to the among herd variance used in the regression for numbers are found in Table 6. The regression accounted only for the random effect of herds while the adjustment for fixed effects was included in the means on which the group averages were regressed. Since five year averages were used, years were included in the means. Other fixed effects were four seasons, three sexes and two managements. Thus, 24 means were calculated for each year in each breed. Figures 1 and 2 illustrate the effects of the adjustment of group averages on the rankings of individuals from different groups.

Table 6. Hierarchal analyses of variance of weaning weights and weaning grades

Source	df	MS	Component	Percentage
Weaning weight				
Subclass	738	4717815.66		37.36
Herds/subclass	2457	16258.72	1652.18	23.01
Within herds	25343	2846.12	2846.12	39.63
Within herd variance/among herd variance = 1.72				
Weaning grade				
Subclass	733	889.00		24.95
Herds/subclass	2348	5.22	0.50	21.57
Within herds	24181	1.24	1.24	53.48
Within herd variance/among herd variance = 2.48				

The number of observations in each of two groups at which the adjusted ratios would be equal if both observations were the same amount above their respective group averages, which were both regressed on the same mean, is given by the formula:

$$N = (1.72)\ \frac{-\mu(x_1 - x_2) - (\bar{x}_2 x_1 - \bar{x}_1 x_2)}{2(\bar{x}_2 x_1 - \bar{x}_1 x_2)}$$

$$\pm\ \frac{[\mu(x_1 - x_2) + (\bar{x}_2 x_1 - \bar{x}_1 x_2)]^2 - 4\mu(x_1 - x_2)(\bar{x}_2 x_1 - \bar{x}_1 x_2)}{2(\bar{x}_2 x_1 - \bar{x}_1 x_2)}$$

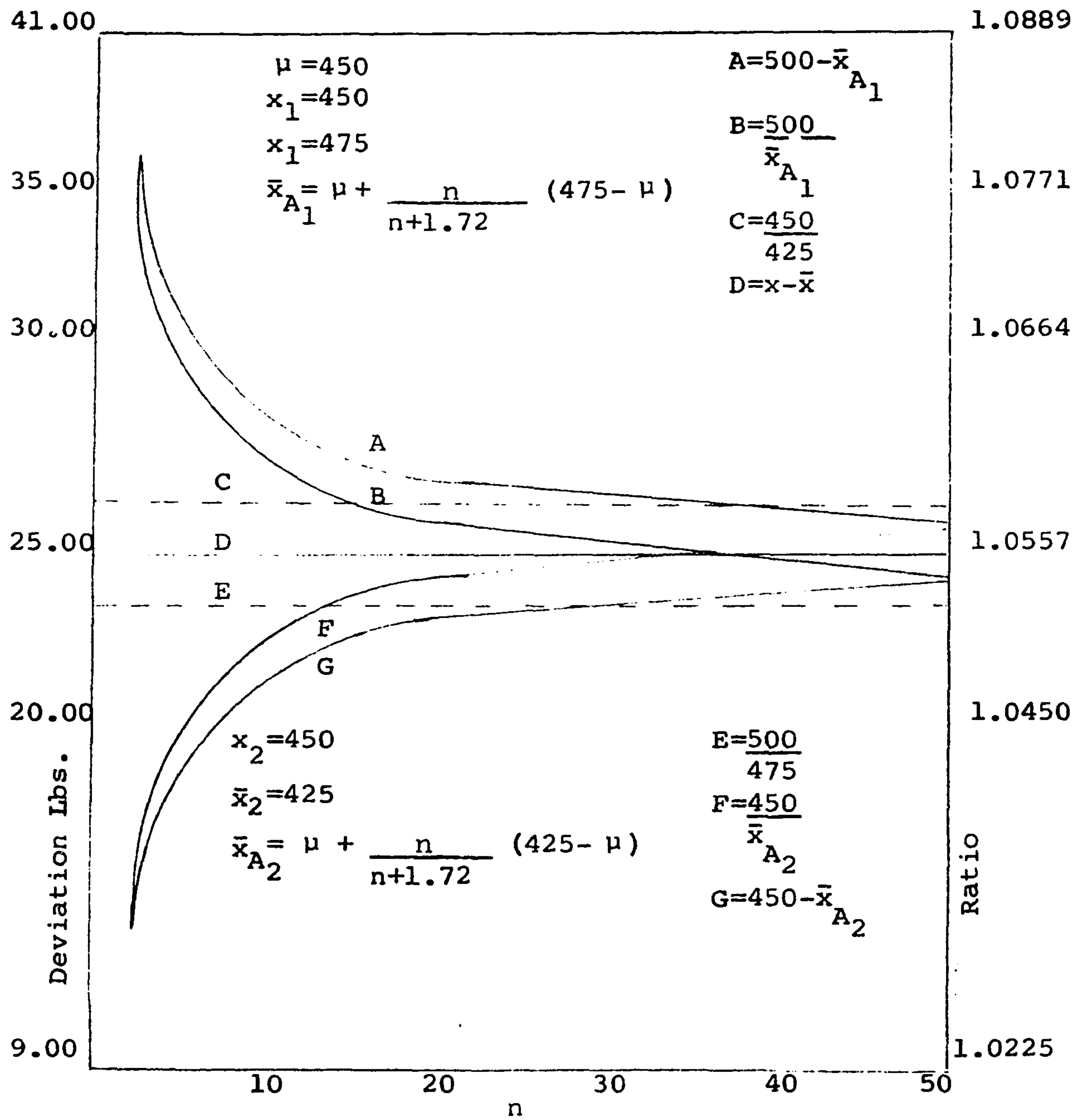

Figure 1. A comparison of the effects of the adjustment for numbers on two observations having the same deviation above different group averages when expressed as ratios and as deviations

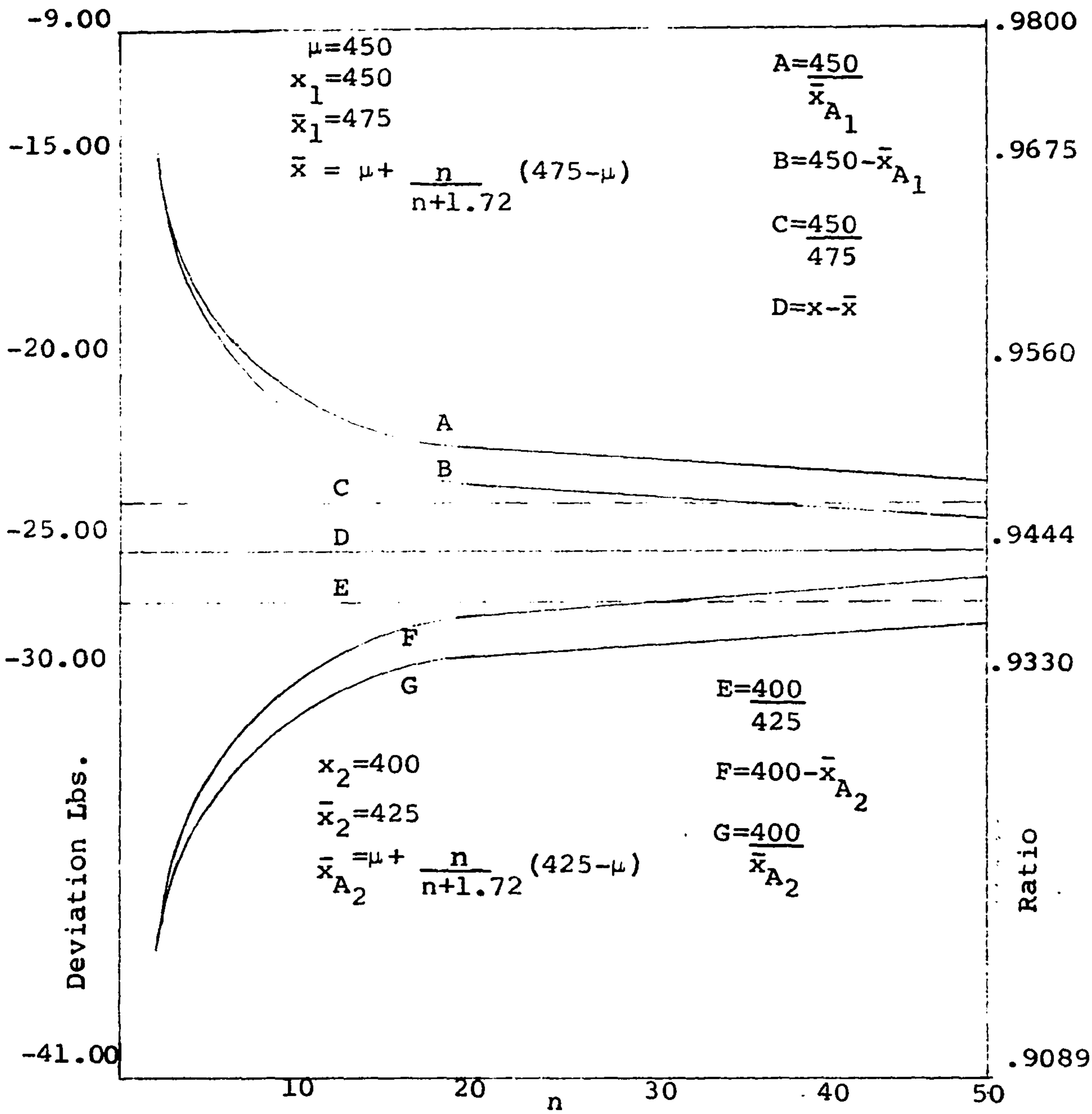

Figure 2. A comparison of the effects of the adjustment for numbers on two observations having the same deviation below different group averages when expressed as ratios and as deviations

The ratio of an observation, whose denominator is not adjusted for the number of observations, that is a fixed amount above the group average will have a larger ratio if the average is small. Therefore, observations above the group average will rank higher from groups with low averages than from groups with high averages if both groups are members of the same population and have equal variances. When both observations are above the group averages by the same amount, the effect of the adjustment for small numbers is to lower the group averages that are above the mean and raise those averages below the mean such that the observation from the group that had the highest average also has the highest ratio. Therefore, the adjustment reverses the rankings of these individuals when numbers are small.

In the example shown in Figure 1, the two observations would rank the same when the number in each group equals 31. The observation from the group with the high average would have the higher ratio of the two when group numbers were less than 31 and the lower ratio of the two when group numbers were greater than 31.

Rankings are not changed for individuals below the group averages in different groups, but relative differences change as shown in Figure 2.

The graphs in Figures 1 and 2 are appropriate for observations of weaning weight and would change according to the appropriate regression. The regression for weaning grade would be $\frac{n}{n + 2.48}$.

<u>Examination of group means</u>

Hierarchal analyses of variance were used to see if expressing the data as ratios and deviations eliminated group effects. Neither ratios nor deviations had sizable sums of squares for any factor and all percentages of variance accounted for by the factors were estimated to be zero. Subsequent examination of expected mean squares revealed that all corrected sums of squares for the factors had expected values of zero.

Ratios and deviations adjusted for group numbers had larger sums of squares than unadjusted data, but in no case were the effects of a factor large enough to account for a significant portion of the total variation.

<u>Examination of variances</u>

The results of the least squares analyses where the variables were the group variances of the adjusted and unadjusted ratios and deviations are summarized in Tables 7 and 8. Regressions of group averages for small numbers had only minor effects on the variances. Main effects for breeds were ignored since the breed with the greatest variance (Red Angus) made up only 3 of the 2,478 groups in this study. In addition, comparisons are normally made

Table 7. Least squares analyses of variance

Source of variation	df	MS	F-value	% Variance
Variance of weaning weight ratios				
Breed	5	0.001481	6.699***	1.32
Season	3	0.002577	11.656***	1.65
Sex	2	0.000094	0.424	-
Management	1	0.000266	1.203	-
Breed x sex	9	0.000275	1.246	0.16
Breed x management	4	0.000388	1.753	0.32
Season x sex	6	0.000436	1.972	0.45
Season x management	3	0.000083	0.377	-
Sex x management	2	0.000303	1.373	0.08
Remainder	2442	0.000221		96.00
Variance of adjusted weaning weight ratios				
Breed	5	0.001532	7.834***	1.58
Season	3	0.002484	12.700***	1.80
Sex	2	0.000102	0.519	-
Management	1	0.000279	1.429	0.03
Breed x sex	9	0.000261	1.333	0.21
Breed x management	4	0.000325	1.663	0.28
Season x sex	6	0.000365	1.866	0.40
Season x management	3	0.000016	0.084	-
Sex x management	2	0.000207	1.059	0.01
Remainder	2442	0.000196		95.69

***P < .005.

Table 7. (Continued)

Source of variation	df	MS	F-value	% Variance
Variance of weaning grade ratios				
Breed	5	0.000992	4.490***	0.82
Season	3	0.000076	0.345	-
Sex	2	0.000139	0.628	-
Management	1	0.003107	14.063***	1.02
Breed x sex	9	0.000116	0.524	-
Breed x management	4	0.000762	3.448*	1.05
Season x sex	6	0.000240	1.087	0.04
Season x management	3	0.000078	0.353	-
Sex x management	2	0.000076	0.342	-
Remainder	2442	0.0000221		97.06
Variance of adjusted weaning grade ratios				
Breed	5	0.001008	4.597***	0.84
Season	3	0.000070	0.320	-
Sex	2	0.000132	0.601	-
Management	1	0.003148	14.362***	1.04
Breed x sex	9	0.000112	0.510	-
Breed x management	4	0.000760	3.469**	1.63
Season x sex	6	0.000232	1.057	0.03
Season x management	3	0.000091	0.416	-
Sex x management	2	0.000070	0.320	-
Remainder	2442	0.000219		96.46

*P < .05.
**P < .01.

Table 7. (Continued)

Source of variation	df	MS	F-value	% Variance
Variance of weaning weight deviations				
Breed	5	83780694.05	11.861***	2.47
Season	3	54864572.32	7.767***	1.03
Sex	2	22013101.67	3.116*	0.24
Management	1	26404045.43	3.738	0.21
Breed x sex	9	11928672.36	1.689	0.44
Breed x management	4	33107164.87	4.687***	1.54
Season x sex	6	9615271.34	1.361	0.16
Season x management	3	1202511.05	0.170	-
Sex x management	2	2778761.87	0.393	-
Remainder	2442	7063734.47		93.91
Variance of adjusted weaning weight deviations				
Breed	5	83798631.52	11.986***	2.50
Season	3	53573181.58	7.663***	1.01
Sex	2	22064183.61	3.156*	0.24
Management	1	27726815.84	3.966*	0.23
Breed x sex	9	12021703.67	1.719	0.46
Breed x management	4	33102944.36	4.735***	1.56
Season x sex	6	9971032.53	1.426	0.19
Season x management	3	1428734.02	0.204	-
Sex x management	2	2852627.63	0.408	-
Remainder	2442	6991471.11		93.81

Table 7. (Continued)

Source of variation	df	MS	F-value	% Variance
Variance of weaning grade deviations				
Breed	5	32.584	2.217*	0.29
Season	3	2.993	0.204	-
Sex	2	8.145	0.554	-
Management	1	77.386	5.265**	0.34
Breed x sex	9	8.912	0.606	-
Breed x management	4	23.516	1.600	0.26
Season x sex	6	9.974	0.679	-
Season x management	3	7.989	0.544	-
Sex x management	2	0.886	0.060	-
Remainder	2442	14.698		99.10
Variance of adjusted weaning grade deviations				
Breed	5	24.700	2.966*	0.47
Season	3	1.423	0.171	-
Sex	2	6.673	0.801	-
Management	1	57.706	6.929**	0.47
Breed x sex	9	3.540	0.425	-
Breed x management	4	15.990	1.920	0.40
Season x sex	6	8.348	1.002	-
Season x management	3	2.843	0.341	-
Sex x management	2	1.179	0.142	-
Remainder	2442	8.328		98.65

Table 8. Least squares constants and means for significant sources of variation in among group variances

Source of variation	Least squares constant	Least squares mean
Variance of weaning weight deviations		
Season		
Winter	- 559.7611	2796.8859
Spring	- 175.1752	3181.4718
Summer	446.9366	3803.5837
Fall	287.9997	3644.6468
Sex		
Bulls	688.9368	4045.5839
Steers	36.5703	3393.2174
Heifers	- 725.5071	2631.1400
Management		
Creep-fed	195.2725	3551.9195
Noncreep-fed	- 195.2725	3161.3746
Variance of weaning grade deviations		
Management		
Creep-fed	- 0.3343	1.1726
Noncreep-fed	0.3343	1.8412
Variance of weaning weight ratios		
Season		
Winter	- 0.00348	0.01133
Spring	0.00141	0.01340
Summer	0.00292	0.01773
Fall	0.00197	0.01678

Table 8. (Continued)

Source of variation	Least squares constant	Least squares mean
Variance of weaning grade ratios		
Management		
Creep-fed	- 0.002118	0.006046
Noncreep-fed	0.002118	0.010283

only within breeds. Ignoring breed effects, the variances of weight ratios were significantly different only for seasons. Variances of grade ratios were significantly different for management (P < .005) and breed by management interaction (P < .01). The group variances of weaning weight deviations were significantly different for season (P < .005), sex (P < .05) and breed by management (P < .005). The variances of weaning grade deviations were significantly different (P < .01) only for management groups.

The percentages of variance accounted for by the effects in the model as given in Table 7 were small even though the effects were highly significant. Only the main effects accounted for more than 1 percent of the variance in weaning weight ratios. The breed by management interaction accounted for more than 1 percent of the variance in weaning grade ratios while for deviations the breed by management interaction was over 1 percent only for weight. Really very little of the total variance was defined in any of the analyses.

Least squares means and constants for the significant causes of differences among group variances are found in Table 8. Again, main effects and interactions for breeds are omitted.

Use of ratios and deviations in genetic analyses

Expected mean squares   Van Vleck, et al. (1961) examined the variance of deviations from herd-year-season averages for dairy records. Their work provided the basis for much of the following examination of genetic uses of ratios and deviations. A brief review of the method of determining meaningful variance components from deviated records follows.

The model for an observation was

$$y_{ijk} = \mu + g_i + s_j + e_{ijk}$$

where

y_{ijk} = the effect of the trait for the k^{th} offspring by the j^{th} sire in the i^{th} group,

μ = effect due to the mean,

g_i = effect due to the i^{th} group,

s_j = effect due to the j^{th} sire, and

e_{ijk} = random error associated with the k^{th} individual by the j^{th} sire and in the i^{th} group.

A deviation was

$$d_{ijk} = (y_{ijk} - \bar{y}_{i..})$$

where

$$\bar{y}_{i..} = \sum_{jk} (\mu + g_i + s_j + e_{ijk}) \frac{1}{n_{i..}} .$$

$$\mathrm{Var}(y_{ijk} - \bar{y}_{i..}) = \mathrm{Var}(d_{ijk}) = E[(d_{ijk}) - \frac{E(d_{ijk})}{n}]^2 .$$

Then,

$$E(d_{ijk}) = E[\mu + g_i + s_j + e_{ijk} - \mu - g_i - \frac{\sum_j n_{ij.} s_j}{n_{i..}}$$

$$- \frac{\sum_j \sum_k e_{ijk}}{n_{i..}}] \; .$$

And since s_j and e_{ijk} are considered to be random variables, $E(s_j) = E(e_{ijk}) = 0$, and $E(d_{ijk}) = 0$. So, the $Var(d_{ijk}) = E(d_{ijk}^2)$.

Let $d_{ij'k'} = k'^{th}$ particular individual by the j' sire.

$$E(d_{ij'k'})^2 = E(y_{ij'k'} - \bar{y}_{i..})^2$$

$$= E[s_j' - \frac{\sum_j n_{ij.} s_j}{n_{i..}} + e_{ij}'k' - \frac{\sum_{jk} e_{ijk}}{n_{i..}}]^2$$

We assume

$$\underset{j \neq j'}{E}(s_j, s_j) = E(s_j, e_{ijk}) = \underset{k \neq k'}{E}(e_{ijk}, e_{ijk'}) = 0 \; .$$

Then,

$$E(d_{ij'k'})^2 = E[s_{j'}^2 - \frac{2 \sum_j n_{ij.}}{n_{i..}} s_j s_{j'} + \frac{\sum_j n_{ij.}^2}{n_{i..}^2} s_j^2 + e_{ij'k'}^2$$

$$- \frac{2e_{ij'k'}^2}{n_{i..}} + \frac{\sum_{jk} e_{ijk}^2}{n_{i..}^2}]$$

$$= \sigma_s^2 - \frac{2n_{ij'.}\ \sigma_s^2}{n_{i..}} + \frac{\sum_j n_{ij.}^2\ \sigma_s^2}{n_{i..}^2} + \sigma_e^2 - \frac{2\sigma_e^2}{n_{i..}} + \frac{n_{i..}}{n_{i..}^2}\ \sigma_e^2$$

$$= \sigma_s^2 - \frac{2n_{ij'.}\ \sigma_s^2}{n_{i..}} + \frac{\sum_j n_{ij.}^2\ \sigma_s^2}{n_{i..}^2} + \sigma_e^2 - \frac{\sigma_e^2}{n_{i..}} .$$

Further simplification was obtained by Van Vleck, et al. (1961) by assuming that all individuals were unrelated. For this special case $n_{ij.} = n_{ij'.} = 1$ and $\frac{\sum_j n_{ij.}^2}{n_{i..}^2} = \frac{n_{i..}}{n_{i..}^2}$, and the $\mathrm{Var}(d_{ij'k'}) = \frac{(n_{i..} - 1)}{n_{i..}} (\sigma_s^2 + \sigma_e^2)$.

The normal structure of beef cattle data is such that this assumption is rarely met. It is, however, apparent that mean and group effects are removed and the among groups effects are reasonably estimated to be zero. The general formula for the variance of a deviation may be simplified when all sires have an equal number of offspring in a group. Then,

$$\frac{\sum_j n_{ij.}^2}{n_{i..}^2} = \frac{n_{ij'.}}{n_{i..}}$$

and

$$\mathrm{Var}(d_{ij'k'}) = \frac{1 - n_{ij'.}}{n_{i..}}\ \sigma_s^2 + \frac{n_{i..} - 1}{n_{i..}}\ \sigma_e^2$$

where

$n_{ij.}'$ = the number of offspring by a sire, and

$n_{i..}$ = the number of individuals in a group.

Expected values for ratios are somewhat more difficult since the concept of expected values has not been shown for nonlinear forms. First efforts to find expected values for mean squares of ratios led only to summations of nonlinear forms and no satisfactory method of estimating variance components.

The data for ratios can be expressed in the following manner.

$$y_{ijkl} = \mu + h_i + g_{ij} + s_{ijk} + e_{ijkl}$$

where

y_{ijkl} = an observation on the l^{th} progeny by the k^{th} sire, in the j^{th} group, in the i^{th} herd,

μ = effect due to the mean,

h_i = effect due to the i^{th} herd,

g_{ij} = effect due to the j^{th} group in the i^{th} herd,

s_{ijk} = effect due to the k^{th} sire in the j^{th} group and the i^{th} herd, and

e_{ijkl} = random error.

An observation expressed as a ratio was

$$r_{ijkl} = \frac{y_{ijkl}}{\bar{y}_{ij..}},$$

where

$$\bar{y}_{ij..} = \frac{1}{n_{ij..}} \sum_{kl} (\mu + h_i + g_{ij} + s_{ijk} + e_{ijkl}) .$$

By assuming that $\bar{y}_{ij..}$ was measured without error or that $\bar{y}_{ij..} = \mu + h_i + g_{ij}$, it was possible to express r_{ijkl} as a weighted linear form and to find expected mean squares consistent with the values obtained in analysis of variance procedures.

Now,

$$r_{ijkl} = \frac{\mu + h_i + g_{ij} + s_{ijk} + e_{ijkl}}{\mu + h_i + g_{ij}}$$

$$= 1 + \frac{s_{ijk} + e_{ijkl}}{\mu + h_i + g_{ij}}$$

and, since μ is a constant and h_i and g_{ij} are random values, $r_{ijkl} = 1 + s_{ijk'} + e_{ijkl'}$.

This model leads to rather straightforward estimates of expected mean squares, and demonstrates that the uncorrected sum of squares for herds, groups and the correction factor (sum of squares due to the mean) is each equal to the total number of observations. This may be shown as follows:

$$\text{Variance } (r_{ijkl}) = \text{Var}(1 + s'_{ijk} + e'_{ijkl})$$

$$= E\left[(1 + s'_{ijk} + e'_{ijkl}) - E(1 + s'_{ijk} + e'_{ijkl}) \right]^2 .$$

Since s'_{ijk} and e'_{ijkl} are random variables with expected values of zero, $E(1 + s'_{ijk} + e'_{ijkl}) = 1$.

Therefore, the variance of the ratio equals

$$E\left[1 + s'_{ijk} + e'_{ijkl} - 1\right]^2 = E(s'_{ijk} + e'_{ijkl})^2$$

assuming the covariance $(s'_{ijk}, s'_{ijk'}) = \mathrm{Cov}(s'_{ijk}, e'_{ijkl})$ $= \mathrm{Cov}(e'_{ijkl}, e'_{ijkl'}) = 0$, and $E(s'_{ijk})^2 = \sigma^2_{s'}$, and $E(e'_{ijkl})^2 = \sigma^2_{e'}$, so $E(s'_{ijk} + e'_{ijkl})^2 = \sigma^2_{s'} + \sigma^2_{e'}$.

The expected mean squares for an among herds, groups within herds, sires within groups and herds and within sire analysis were found as follows:

Uncorrected total sum of squares (UTSS)

$$E\,[UTSS] = E \sum_{ijkl} (r_{ijkl})^2 = \sum_{ijkl} E\left[1 + s'_{ijk} + e'_{ij}\right]^2$$

$$= \sum_{ijkl} (1 + \sigma^2_{s'} + \sigma^2_{e'})$$

$$= n.... + n....\ \sigma^2_{s'} + n....\ \sigma^2_{e'}$$

Uncorrected sire sum of squares (USSS)

$$E\,[USSS] = E \sum_{ijk} \frac{(r^2_{ijk.})}{n_{ijk.}} = E \sum_{ijk} \frac{1}{n_{ijk.}} \sum_{l} (1 + s'_{ijk} + e'_{ijkl})^2$$

$$= \sum_{ijk} \frac{1}{n_{ijk.}} E\left[n_{ijk.} + n_{ijk.}\ s'_{ijk} + \sum_{l} e'_{ijkl}\right]^2$$

$$= \sum_{ijk} n_{ijk.} + n_{ijk.}\ \sigma_s^2 + \frac{\sum_l \sigma_e^2}{n_{ijk.}}$$

$$= n_{....} + n_{....}\ \sigma_s^2 + \sum_{ijk} \sigma_e^2 .$$

Correction factor (C.F.)

$$E[C.F.] = E\ \frac{r^2_{....}}{n_{....}} = E\ \frac{1}{n_{....}} \left[\sum_{ijkl} (1 + s'_{ijk} + e'_{ijkl}) \right]^2$$

Since the ratios by definition must average one when averaged for each group, and $s'_{ijk} + e'_{ijkl}$ are deviations from the group average and when summed within the group must sum to zero,

$$\sum_{kl} (s'_{ijk} + e'_{ijkl}) = 0 .$$

Then,

$$\frac{1}{n_{....}} \sum_{ij} \sum_{kl} (1 + s'_{ijk} + e'_{ijkl})^2 = \frac{1}{n_{....}} \left[\sum_{ij} n_{ij..} \right]^2$$

$$= \frac{n^2_{....}}{n_{....}} = n_{....}$$

Uncorrected group sum of squares (UGSS)

$$E[UGSS] = E \sum_{ij} \frac{r^2_{ij..}}{n_{ij..}} = E \sum_{ij} \frac{1}{n_{ij..}} \sum_{kl} (1 + s'_{ijk} + e'_{ijkl})^2$$

$$= \sum_{ij} \frac{n^2_{ij..}}{n_{ij..}} = n_{....}$$

Uncorrected herd sum of squares (UHSS)

$$E[\text{UHSS}] = E \sum_i \frac{r_{i\ldots}^2}{n_{i\ldots}} = E \sum_i \frac{1}{n_{i\ldots}} \sum_j \sum_{kl} (1 + s'_{ijk} + e'_{ijkl})^2$$

$$= \frac{\sum_i n_{i\ldots}^2}{n_{i\ldots}} = n_{\ldots.}$$

These expected values show that ordinary expected values for mean squares are not likely to be appropriate for explaining the variance observed in ratios of beef cattle records to their group means. However, this model fails to explain the loss of genetic variation that occurs when sires are confounded with groups. Therefore other models were examined.

A deviation model was chosen because of the success of Van Vleck, _et al_. (1961) in explaining observed variances in dairy cattle deviation records.

The model was

$$y_{ijkl} = \mu + h_i + g_{ij} + s_{ijk} + e_{ijkl},$$

where the model was previously defined.

$$r_{ijkl} = \frac{\mu + h_i + g_{ij} + s_{ijk} + e_{ijkl}}{\mu + h_i + g_{ij}},$$

Let $\bar{s}_{ij}$ = average sire effect in a group, and

$\bar{e}_{ij}$ = average environmental effect in a group.

Then,

$$r_{ijkl} = \frac{\mu + h_i + g_{ij} + \bar{s}_{ij} + \bar{e}_{ij} + s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij}}{\mu + h_i + g_{ij} + \bar{s}_{ij} + \bar{e}_{ij}}$$

$$= 1 + \frac{s_{ijk} - \bar{s}_{ij}}{\mu + h_i + g_{ij} + \bar{s}_{ij} + \bar{e}_{ij}}$$

$$+ \frac{e_{ijkl} - \bar{e}_{ij}}{\mu + h_i + g_{ij} + \bar{s}_{ij} + \bar{e}_{ij}}$$

$$= 1 + (s_{ijk} - \bar{s}_{ij})' + (e_{ijkl} - \bar{e}_{ij})'.$$

Since sire and error effects are considered to be random variables with expected values of zero, $E[1 + (s_{ijk} - \bar{s}_{ij})' + (e_{ijkl} - \bar{e}_{ij})'] = 1$. Therefore, using the prime to indicate a particular observation, e.g., the l^{th} offspring from the k^{th} sire,

$$\begin{aligned} \mathrm{Var}(r_{ijk'l'}) &= E[1 + s_{ijk'} - \bar{s}_{ij} + e_{ijkl} - \bar{e} - E(1 + s_{ijk'} \\ &\qquad - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij})]^2 \\ &= E[s_{ijk'} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij}]^2 \end{aligned}$$

Assuming

$$\begin{aligned} \mathrm{Cov}(s_{ijk}, s_{ijk'}) &= \mathrm{Cov}(s_{ijk}, e_{ijkl}) \\ &= \mathrm{Cov}(e_{ijkl}, e_{ijkl'}) = 0 \end{aligned}$$

$$= E\left[s^2_{ijk'} - \frac{2 \sum_k n_{ijk.} \, s_{ijk'} \, s_{ijk}}{n_{ij..}} + \frac{\sum_k n^2_{ijk.}}{n^2_{ij..}} s^2_{ijk} \right.$$

$$+ e^2_{ijk'l'} - \frac{2 \sum_{kl} e_{ijkl} \, e_{ijk'l'}}{n_{ij..}}$$

$$\left. + \frac{\sum_{kl} e^2_{ijkl}}{n^2_{ij..}} \right]$$

$$= \sigma^2_s - \frac{2n_{ijk.'} \, \sigma^2_s}{n_{ij..}} + \frac{\sum_k n^2_{ijk.} \, \sigma^2_s}{n^2_{ij..}} + \sigma^2_e - \frac{\sigma^2_e}{n_{ij..}}$$

Thus, the variances for deviations and ratios were found to contain the same components except that variances for ratios were weighted by the squared group means.

Expected mean squares were then calculated for ratios from this model. These expected mean squares should be directly applicable to deviations by simply omitting the effect due to the mean and the weighting of the variance by the square of the group mean.

Expected mean squares for the model with sires nested within groups are found as follows:

$$r_{ijkl} = 1 + s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij}$$

$$E[UTSS] = E \sum_{ijkl} r^2_{ijkl}$$

$$= E \sum_{ijkl} [1 + s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij}]^2$$

and assuming that covariances between sires and errors are zero,

$$= E\left[\sum_{ijkl} 1 + s^2_{ijk'} - \frac{2n_{ijk.'}}{n_{ij..}} s^2_{ijk'} + \frac{\sum_k n^2_{ijk.}}{n^2_{ij..}} s^2_{ijk} + e^2_{ijkl} - \frac{2e^2_{ijk'l'}}{n_{ij..}} + \frac{\sum_{kl} e^2_{ijkl}}{n^2_{ij..}}\right]$$

$$= \sum_{ijkl}\left[1 + \sigma^2_s - \frac{2n_{ijk.'}}{n_{ij..}} \sigma^2_s + \frac{\sum_k n^2_{ijk.}}{n^2_{ij..}} \sigma^2_s + \sigma^2_e - \frac{2\sigma^2_e}{n_{ij..}} + \frac{\sigma^2_e}{n_{ij..}}\right]$$

$$= n_{....} + \left(n_{....} - \frac{2 \sum_{ijk} n^2_{ijk.}}{n_{ij..}} + \frac{\sum_{ijk}\sum_k n^3_{ijk.}}{n^2_{ij..}}\right) \sigma^2_s + \left(n_{....} - \sum_{ij} n_{ij}\right) \sigma^2_e$$

$$E[USSS] = E \sum_{ijk} \frac{r^2_{ijk.}}{n_{ijk.}} = E \sum_{ijk} \frac{1}{n_{ijk.}} \left[n_{ijk.} + n_{ijk.} (s_{ijk} - \bar{s}_{ij}) + \sum_l (e_{ijkl} - \bar{e}_{ij})\right]^2$$

$$= E\left[\sum_{ijk} \frac{1}{n_{ijk.}}\left[n^2_{ijk.} + n^2_{ijk.}\,(s_{ijk} - \bar{s}_{ij})^2\right.\right.$$

$$\left.\left. + \sum_{l}(e_{ijkl} - \bar{e}_{ij})^2\right]\right.$$

$$= \sum_{ijk} E\left[n_{ijk.} + n_{ijk.}\,(s_{ijk} - \bar{s}_{ij})^2 + (e^2_{ijkl}\right.$$

$$\left. - \frac{2e^2_{ijkl}}{n_{ij..}} + \frac{e^2_{ijkl}}{n_{ij..}})\right]$$

$$= n_{....} + n_{....}\,\sigma^2_s - 2\sum_{ijk}\frac{n^2_{ijk.}\,\sigma^2_s}{n_{ij..}}$$

$$+ \sum_{ijk}\frac{(\sum_k n^3_{ijk.})\,\sigma^2_s}{n^2_{ij..}} + \sigma^2_e\,(\sum_{ijk} n_{ijk} - \sum_{ij} n_{ij})$$

$$E[C.F.] = \frac{E\,r^2_{....}}{n_{....}}$$

$$= \frac{1}{n_{....}} E \sum_{ijkl}(1 + s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij})^2$$

$$\bar{s}_{ij} = \frac{\sum_{kl} s_{ijk}}{n_{ij..}} \quad \text{and} \quad \bar{e}_{ij} = \frac{\sum_{kl} e_{ijkl}}{n_{ij..}}$$

Therefore, $\sum_{kl}(s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij})$ = the sum of deviations about their mean and must equal zero.

$$\text{Now, } E[C.F.] = \frac{1}{n_{....}} \sum_{ijkl} (1)^2$$

$$= \frac{n^2_{....}}{n_{....}} = n_{....}$$

$$E[UGSS] = E \sum_{ij} \frac{r^2_{ij..}}{n_{ij..}} = E \sum_{ij} \frac{1}{n_{ij..}} \Big[\sum_{kl} (1 + s_{ijk} - s_{ij} + e_{ijkl} - \bar{e}_{ij}) \Big]^2$$

$$= \sum_{ij} \frac{n^2_{ij..}}{n_{ij..}} = n_{....}$$

$$E[UHSS] = E \sum_{i} \frac{r^2_{i...}}{n_{i...}} = E \sum_{i} \frac{1}{n_{i...}} \Big[\sum_{j} \sum_{kl} (1 + s_{ijk} - \bar{s}_{ij} + e_{ijkl} - \bar{e}_{ij}) \Big]^2$$

$$= \sum_{i} \frac{n^2_{i...}}{n_{i...}} = n_{....}$$

Examination of a model where sires were present in more than one group led to very similar results. The model and expected mean squares for the uncorrected total sums of squares and uncorrected sire sums of squares will be included. Herd sums of squares and the correction factor once again equaled the total number of observations and can be demonstrated in the same fashion. The group sum of squares is not easily demonstrated for this model and appears to have little meaning. Therefore, it will not be included.

The model was:

$$y_{ijkl} = \mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i + s_{ik} + \bar{s}_i + e_{ijkl} - \bar{e}_i$$

where

$$\bar{s}_i = \sum_k \frac{n_{i.k.} s_{ik}}{n_{i...}} ,$$

$$\bar{e}_i = \sum_{jkl} \frac{e_{ijkl}}{n_{i...}} ,$$

s_{ik} = an effect due to the k^{th} sire in the i^{th} herd, and

all other elements of the model are as previously described.

$$r_{ijkl} = \frac{\mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i + s_{ik} - \bar{s}_i + e_{ijkl} - \bar{e}_i}{\mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i}$$

$$= 1 + s_{ik} - \bar{s}_i + e_{ijkl} - \bar{e}_i .$$

$$E[UTSS] = E \sum_{ijkl} r^2_{ijkl} = E \sum_{ijkl} [1 + s_{ik} - \bar{s}_i$$

$$+ e_{ijkl} - \bar{e}_i]^2$$

Assuming $Cov(s_{ik}, e_{ijkl}) = Cov(\bar{s}_i, e_{ijkl}) = Cov(e_{ijkl}, e_{ijkl'}) = 0$

$$E[UTSS] = E \sum_{ijkl} [1 + (s_{ik} - \bar{s}_i)^2 + (e_{ijkl} - \bar{e}_i)^2]$$

$$= E\ [n_{....} + \sum_{ijkl} (s^2_{ik'} - \frac{2n_{i.k.'}s_{ik}s_{ik'}}{n_{i...}}$$

$$+ \frac{\sum_k n^2_{i.k.} s^2_{ik}}{n^2_{i...}}) + \sum_{ijkl} (e^2_{ijkl} - \frac{2e^2_{ijkl}}{n_{i...}}$$

$$+ \frac{\sum_{jkl} e^2_{ijkl}}{n^2_{i...}})]$$

$$= n_{....} + n_{....}\ \sigma^2_s - 2 \frac{\sum_{ijkl} n_{i.k.'}}{n_{i...}} \sigma^2_s$$

$$+ \sum_{ijkl} \frac{(\sum_k n^2_{i.k.})}{n^2_{i...}} \sigma^2_s + n_{....}\ \sigma^2_e$$

$$- 2 \sum_i \sigma^2_e + \sum_i \sigma^2_e$$

$$= n_{....} + (n_{....} - 2 \sum_{ik} \frac{n^2_{i.k.}}{n_{i...}}$$

$$+ \sum_{ik} \frac{(\sum_k n^3_{i.k.})\ \sigma^2_s}{n^2_{i...}} + (n_{...} - \sum_i n_i)\ \sigma^2_e$$

$$E[USSS] = E \sum_{ik} \frac{r_{i.k.}^2}{n_{i.k.}}$$

$$= E \sum_{ik} \frac{1}{n_{i.k.}} \left[\sum_{jl} (1 + s_{ik} - \bar{s}_i + e_{ijkl} - \bar{e}_i) \right]^2$$

$$= E \sum_{ik} \frac{1}{n_{i.k.}} \left[n_{i.k.} + n_{.k.} (s_{ik} - \bar{s}_i) \right.$$

$$\left. + \sum_{jl} (e_{ijkl} - \bar{e}_i) \right]^2$$

If we assume $\mathrm{Cov}(s_{ik}, e_{ijkl}) = \mathrm{Cov}(s_{ik}, s_{ik'}) =$ $\mathrm{Cov}(e_{ijkl}, e_{ijkl'}) = 0$, then

$$E[USSS] = E\left[\sum_{ik} \frac{1}{n_{i.k.}} \left(n_{i.k.}^2 + n_{i.k.}^2 (s_{ik} - \bar{s}_i)^2 \right.\right.$$

$$\left.\left. + \sum_{jl} \frac{(e_{ijkl} - \bar{e}_i)^2}{n_{i.k.}} \right) \right]$$

$$= \sum_{ik} \quad n_{i.k.} + n_{i.k.} \left(\sigma_s^2 - \frac{2n_{i.k.} \; \sigma_s^2}{n_{i...}} \right.$$

$$\left. + \frac{\sum_k n_{i.k.}^2 \; \sigma_s^2}{n_{i...}^2} \right) + \left(\sigma_e^2 - \frac{\sigma_e^2}{n_{i...}} \right)$$

$$= n.... + (n.... - 2 \; \frac{\sum_{ik} n_{i.k.}^2}{n_{i...}}$$

$$+ \sum_{ik} \frac{(\sum_k n_{i.k.}^3)}{n_{i...}^2}) \; \sigma_s^2 + \sum_{ik} \frac{(n_{i...} - 1)}{n_{i...}} \sigma_e^2$$

These models may help to explain how deviations and ratios can remove genetic effects as well as identified sources of environmental variation when sires are partially or completely confounded with groups.

<u>Repeatability and most probable producing ability</u>

Repeatability estimates are frequently used as an important predictive tool. Repeatability will be defined as the correlation among different records for the same trait measured on the same individual, or as the fraction of the total variance accounted for by genotype and permanent environmental effects. Repeatability may be illustrated in the following example:

Let the model be

$$y_{ij} = \mu + c_i + e_{ij}$$

where

y_{ij} = weaning weight of the j^{th} calf from the i^{th} cow,

μ = mean weaning weight,

c_i = effect due to the i^{th} cow = $g_i + e_{pi}$ = genetic effect plus permanent environmental effect due to the i^{th} cow, and

e_{ij} = random error associated with the j^{th} calf from the i^{th} cow.

The variance of $y_{ij} = \sigma_c^2 + \sigma_e^2$ and since $E(c_i) = E(e_{ij}) = 0$, the covariance

$$\text{Cov}(y_{ij},y_{ij'}) = E\,(\mu + c_i + e_{ij} - \mu)(\mu + c_i + e_{ij'} - \mu)$$

and assuming the

$\text{Cov}(c_i,e_{ij}) = \text{Cov}(e_{ij},e_{ij'}) = 0$, the $\text{Cov}(y_{ij},y_{ij'}) = \sigma_c^2$

and the repeatability is

$$\text{repeatability} = \frac{\sigma_g^2 + \sigma_{e_p}^2}{\sigma_t^2} = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{e_t}^2}$$

$$= \rho_{y_{ij},y_{ij'}} .$$

where

$\sigma_{e_p}^2$ = permanent environmental effect, and

$\sigma_{e_t}^2$ = temporary environmental effect.

The repeatability for ratios was estimated in the same manner.

The model was

$$x_{ij} = \mu + g_i + c_j + e_{ijk}$$

where

x_{ij} = effect of the trait for the j^{th} record on the i^{th} individual,

μ = effect due to the mean,

g_i = effect due to the i^{th} group,

c_j = effect due to the j^{th} cow, and

e_{ijk} = random error.

The ratio was $R_{ijk} = \dfrac{x_{ijk}}{\bar{x}_{i..}}$ where $\bar{x}_{i..}$ is assumed to be a constant or measured without error.

The expected value of R_{ijk} is

$$E[R_{ijk}] = E\,\frac{x_{ijk}}{\bar{x}_{i..}} = \frac{1}{\bar{x}_{i..}}\,E\,(\mu + g_i + c_j + e_{ijk})$$

$$= \frac{\mu + g_i}{\bar{x}_{i..}}$$

since c_{ij} and e_{ijk} are random variables with expected values of zero. And the covariance of $R_{ijk}, R_{ijk'}$ is

$$\mathrm{Cov}\left(\frac{x_{ijk}}{\bar{x}_{i..}}, \frac{x_{i'jk'}}{\bar{x}_{i..'}}\right) = E\,\frac{x_{ijk} - (\mu + g_i)}{\bar{x}_{i..}}\;\frac{x_{i'jk'} - (\mu + g_{i'})}{\bar{x}_{i..}}$$

$$= \frac{1}{\bar{x}_{i..}\bar{x}_{i..'}}\,E\left[\mu + g_i + c_j + e_{ijk} - \mu - g_i\right]\left[\mu + g_{i'}\right.$$

$$\left. + c_j + e_{i'jk'} - \mu - g_{i'}\right]$$

$$= \frac{\sigma_c^2}{\bar{x}_{i..}\bar{x}_{i..'}}\,.$$

The variance of R_{ijk} is

$$\mathrm{Var}(R_{ijk}) = E\left[\frac{x_{ijk}}{\bar{x}_{i..}} - E\left[\frac{(x_{ijk})}{\bar{x}_{i..}}\right]\right]^2 = \frac{\sigma_c^2 + \sigma_e^2}{\bar{x}_{i..}^2}$$

And,

$$\text{Repeatability} = \frac{\dfrac{\sigma_c^2}{\bar{x}_{i..}\bar{x}_{i'..}}}{\dfrac{\sigma_c^2 + \sigma_e^2}{\bar{x}_{i..}^2} \quad \dfrac{\sigma_c^2 + \sigma_e^2}{\bar{x}_{i..'}^2}} = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_e^2}$$

$$= \rho_{R_{ijk}, R_{i'jk'}}$$

Most probable producing ability (MPPA) estimates are normally used to predict performance based on the number of records on the individual and the repeatability of the trait. However, the formula normally used to express MPPA includes a regression of the average of n records of an individual on the population average. The regression coefficient is found by minimizing the squared difference between the average performance of an individual and the true performance value for the individual. The regression coefficient is found by solving for a regression coefficient so the squared difference is minimized. For ratios the regression coefficient may be found as follows:

The model was

$$x_{ij} = \mu + c_i + e_{ij}$$

where the model has been previously defined.

The average of n records was estimated as

$$\bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^{n} (\mu + c_i + e_{ij}) = \mu + c_i + \sum_j \frac{e_{ij}}{n}$$

The average of n ratios was estimated as $\frac{\bar{x}_{i.}}{\mu}$, where μ = a five-year-breed average and was considered to be a population parameter. The ratio of the true effect of the cow was defined to be $\frac{c_i}{\mu}$. To minimize the difference between the average ratio and the true ratio the average ratio of the cow was expressed as a deviation from (1.0), the average ratio of all cows, and the following equation solved for k_i such that k_i was a minimum. The expected value of the function $F = E[k_i \left(\frac{\bar{x}_i}{\mu} - \frac{\mu}{\mu}\right) - \frac{c_i}{\mu}]^2$ was expressed in terms of the model as

$$F = E[k_i \left(\frac{\mu + c_i + \sum_j \frac{e_{ij}}{n} - \mu}{\mu}\right) - \frac{c_i}{\mu}]^2$$

$$= k_i^2 \left[\frac{\sigma_c^2 + \frac{\sigma_e^2}{n}}{\mu^2}\right] - 2\, k_i \frac{\sigma_c^2}{\mu^2} + \frac{\sigma_c^2}{\mu^2} .$$

The partial derivative, the k_i of the function was equated to zero and the function solved for k_i, a minimum.

$$\frac{1}{2} \frac{\partial f}{\partial k_i} = k_i \frac{\sigma_c^2 + \frac{\sigma_e^2}{n}}{\mu^2} - \frac{\sigma_c^2}{\mu^2} = 0$$

$$k_i \left[\frac{\sigma_c^2 + \frac{\sigma_e^2}{n}}{\mu^2}\right] = \frac{\sigma_c^2}{\mu^2}$$

$$k_i = \frac{\sigma_c^2}{\sigma_c^2 + \frac{\sigma_e^2}{n}}$$

k_i was expressed in terms of repeatability (r) in the following way:

$$k_i = \frac{\sigma_c^2}{\sigma_c^2 + \frac{\sigma_e^2}{n}} = \frac{\frac{\sigma_c^2}{\sigma_e^2}}{\frac{\sigma_c^2}{\sigma_e^2} + \frac{1}{n}} \quad \text{and since} \quad \frac{\sigma_c^2}{\sigma_e^2} = \frac{r}{1 - r}$$

$$k_i = \frac{\frac{r}{1 - r}}{\frac{r}{1 - r} + \frac{1}{n}} = \frac{r}{r + \frac{1 - r}{n}} = \frac{nr}{1 + (n - 1)r} .$$

Thus, the most probable producing ability (MPPA) can be expressed as

$$\text{MPPA} = 1 + \frac{nr}{1 + (n - 1)r} \left[\frac{\bar{x}_{i.}}{\mu} - 1 \right] .$$

DISCUSSION

Statistical properties of ratios

A ratio of an observation to a group average is a method of expressing data such that the mean and the effects common to the group are replaced by the constant (1). The effects not common to all members of a group are weighted by the group average and their variances are weighted by the square of the group average if the denominator is measured without error. The use of ratios to adjust or transform the data will tend to equalize the variances of groups if these are proportional to their mean such that the coefficients of variation are equal. If the variance within groups is equal or if the mean and variances of the observations are not proportional, then the ratio would not be an appropriate transformation and could make the variances more unlike among groups.

Standard procedures to estimate the variance of ratios lead to identical values to those expected when assuming the denominator is measured without error.

The use of ratios or deviations to remove effects common to a group will also remove any genetic effects that are confounded with groups.

Preliminary analyses

The preliminary least squares analyses were used to establish the relationships among means and variances of groups and to determine the number of observations in combinations of groups for sex, season, and management classes. These relationships were examined to determine whether they could reasonably explain the observed results when variances and coefficients of variation were analyzed by standard least squares analysis procedures, since there appears to be no *a priori* information for using this procedure.

The analyses substantiated the hypothesis that means and variances were proportional such that coefficients of variation were equal for sex and management subclasses when the variable was weaning weight. These results support the idea that means and variances were not independent, or that variances were not equal among groups, and that some procedure for equalizing variances as well as equalizing means would be necessary to fairly compare individuals from different groups.

The small average size (10.3) and large standard deviation (17.7) in group size suggested that subtraction of or division by a group average that is assumed to be measured without error could lead to substantial errors in rankings. This indicated that the group averages should be adjusted for the number of observations to obtain better estimates when numbers were small.

The least squares analysis of sex, season, management, and their interactions for group averages, variances and coefficients of variation indicated large differences among group averages, substantial differences among group variances and few differences among the coefficients of variation for the factors.

Least squares procedures appeared to be useful for analyzing these statistics and the observed results were consistent with *a priori* information on the relationships of mean and variance for growth traits. The presence of fewer differences among the coefficients of variation further suggested that weighting of observations by group means had some merit as a transformation to obtain equality among the variances of weaning weight groups. This property of independence between group means and variances imparted by the transformation would also make observations weighted by the group average comply with the assumption of equal variances necessary for examining the data by standard analysis of variance procedures. However, unless the group averages were assumed to be measured without error, non-linearity has been introduced because the observations were divided by many different group averages.

Since the comparisons to be made concerned the ratio of an observation to a group average, and the group variance of ratios was an algebraic identity with the square of the coefficient of variation, the variance of ratios was considered to be the appropriate variable to study rather than the coefficients of variation.

No real differences in variances for weaning grades were observed. Therefore, no advantage would be gained by adjusting variances of this trait by using ratios. In fact, the only important differences among means for weaning grades were due to management. This indicates that deviations or, within management, the absolute grade would be a more appropriate way of expressing grade records than using ratios.

Adjustment for numbers

Figures 1 and 2 were used to illustrate an important effect of the adjustment of averages for the number of observations. If two animals from different groups are both a fixed amount, say 25 pounds, above their group averages and the averages are different, e.g., 450 and 500 pounds, then the animal from the group within an average of 450 pounds would have a higher ratio since 25/450 is a larger fraction than 25/500. If both groups come from a population where the true mean is 475 pounds, the adjustment would regress the group averages closer to the true mean. The average of 500 would be decreased, to 487.5

pounds, and the average of 450 would be increased, to 462.5 pounds, when only two individuals were in the averages and the effect would be to make the deviation from the group estimated to have a mean of 487.5, equal to 37.5 pounds and the deviation from the other group would be 12.5 pounds. Since 37.5/487.5 is greater than 12.5/462.5, the larger observation now has a larger ratio and the rankings have been reversed. Rankings of ratios for animals whose group averages were regressed for numbers would be very similar to rankings for deviated records when the averages were also regressed. Only when numbers of observations are large enough to indicate that the groups are not members of the same population and the group averages are good estimates of their respective means would the observation from the group with the lower average have the higher ratio. The number of observations in each of two groups at which the adjusted ratios would be the same if both observations had the same absolute difference above their respective group averages, both of which were regressed on the same mean is given by the following quadratic formula:

$$n = (1.72) \quad \frac{-\mu\,(x_1 - x_2) - (\bar{x}_2 x_1 - \bar{x}_1 x_2)}{2\,(\bar{x}_2 x_1 - \bar{x}_1 x_2)}$$

$$\pm \quad \frac{[\mu(x_1 - x_2) + (\bar{x}_2 x_1 - \bar{x}_1 x_2)]^2 - 4\mu(x_1 - x_2)(\bar{x}_2 x_1 - \bar{x}_1 x_2)}{2(\bar{x}_2 x_1 - \bar{x}_1 x_2)}$$

The value of 1.72 is the appropriate weighting factor for the quadratic equation when the variable is weaning weight. The appropriate weighting factor is the ratio σ_w^2/σ_h^2 which appears in the regression coefficient of the adjustment for numbers. The appropriate factor for grades would be 2.48 in these data.

When observations are an equal amount below their respective group averages, no change in rankings occur, although the relative difference in ratios will change.

It is possible through the adjustment for numbers for all individuals in a small group with a high group average to have ratios greater than one. It is also possible for all individuals in a small group with a low average to have ratios less than one.

Another important consideration when using the regression for numbers is that the mean used to divide the observations or to subtract from the observations is no longer an arithmetic average of the observations. Sums of squares for group effects will no longer be zero and may indicate that substantial amounts of variation that were associated with common elements of the group remain in the observations. Since the purpose of expressing the data in this manner was to remove the variation of the common elements of the group, the sums of squares remaining for group elements pose problems in interpretation. In these data, sums of squares for group elements remained small after adjusting the means for numbers and were not considered to be serious.

The regression for numbers is expected to somewhat limit the range of ratios and deviations. The occurrence of a good observation by chance falling into a small group with a low average resulting in an extremely high ratio is now quite unlikely since the average of this group would be regressed upward toward the group mean. However, group variances will not be decreased since the arithmetic average gives minimum variance.

The inclusion of several elements, i.e., year, season, sex, management, etc., in the group necessitates the calculation of many means and suggests this should be done only in large bodies of data where computer facilities are available.

Examination of group means

Either deviations or ratios appeared to effectively remove differences due to group means. However, the examination of the expected mean squares revealed that the expected values for the corrected sums of squares for any factor common to members of a group, i.e., breed, season, sex, management, herd and year, should be zero. Therefore, estimation of components of variance for groups should be negative when estimated by standard procedures. Failure of the corrected sums of squares to be zero for a factor common to members of a group should indicate an effect that was not removed. Use of regression to adjust group averages for numbers resulted in larger sums of squares for all factors for deviated records. This suggests that increased confidence in the group averages due to regression may be offset by failure to remove all effects common to a group from the data. However, all sums of squares for elements of the model using ratios were small, totals were less than 3.0, and showed no serious deviation from the expected values of zero.

Examination of group variances

Analyses where the variable was variances of the 2,478 groups indicated that the use of ratios effectively removed differences among the group variances of weaning weight that were caused by sex, management and breed by management interaction. Differences among the variances of season subclasses were slightly inflated. However, previous analysis of much of this data (Sellers, 1968) indicated that season accounted for only about one percent of the total variation, so this result may not be serious. Effects for breeds were ignored, although the analyses indicated significant differences.

The analysis of group variances for deviated weaning weight records was essentially the same, whether deviated from adjusted or unadjusted averages. The significant differences among variances for season, sex, management and season by management when using deviated records indicate that this method may not give good comparisons across groups, particularly since sex and management are important sources of variation in weaning weight.

The analyses of grade data revealed that management was the only important source contributing to differences in group variances for deviated records, with creep-fed calves being less variable than noncreep-fed calves. However, the variances for grade ratios were significantly different for breed, management, and breed by management. There appears to be no advantage in using ratios to express grades. But, if comparisons are made on a within breed basis, ratios should only be slightly less effective than deviations for removing effects due to common elements of the group. The larger F-values associated with management for ratios were to be expected since the class with the larger mean (creep-fed) had the lower variance.

Genetic analyses

Considerable difficulty was encountered when attempting to extend the use of deviated records and ratios to genetic analyses. The initial treatment of ratios as nonlinear forms further complicated matters. The primary assumption that the denominator was measured without error allowed reparameterization of the model and the ratio could then be treated as a weighted linear form. Although such an assumption is approached with some trepidation, the observed values for sums of squares associated with herds, groups and group mean are consistent with the expected values found in the reparameterized model. Since the common assumption used to solve systems of equations, that the sums of random effects equal zero, is simply applied to groups, the assumption that the denominator is measured without error may be acceptable in analysis of variance procedures for investigating ratios. Also, the regression for numbers may lend some confidence to estimates of group means.

The model

$$r_{ijkl} = \frac{\mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i + (s_{ik} - \bar{s}_i) + (e_{ijkl} - \bar{e}_i)}{\mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i}$$

$$= 1 + \frac{(s_{ik} - \bar{s}_i) + (e_{ijkl} - \bar{e}_i)}{\mu + h_i + g_{ij} + \bar{s}_i + \bar{e}_i}$$

$$= 1 + (s_{ik} - \bar{s}_i)' + (e_{ijkl} - \bar{e}_i)'$$

in which the sire and environmental effects can be expressed as deviations about some group mean seemed to offer the greatest chance of explaining how removal of group effects also removed genetic variance. The examination of expected mean squares for this model revealed the coefficient

$$[n_{....} - 2 \sum_{ik} \frac{n_{i.k.}^2}{n_{i...}} + \sum_{ik} \sum_{k} \frac{n_{i.k.}^3}{n_{i...}^2}]$$

for the sire component in the among sire corrected sum of squares expectation. Nondeviation models had only the coefficient n...., which showed no loss in genetic variance when sires were confounded with groups. The negative covariance associated with the deviation model subtracts out the sire variance if there is only one sire per group. This could also be shown by summation when only one sire is present per group. However, only the deviated model appears to explain the loss of genetic variance associated with partial confounding of sire and group. If each sire has an approximately equal number of offspring per group and more than one sire is represented, then the coefficient could be used to estimate the loss of sire variance and to adjust for the loss. The proper adjustment would be to multiply the estimate of sire variance found from the analysis of variance by the reciprocal of the coefficient of the sire variance.

These expected values appear to be consistent with those found by Van Vleck, et al. (1961) in the examination of the variance of deviations.

Therefore, the use of either deviations or ratios to estimate genetic components of variance in normal beef cattle data would underestimate the genetic variances due to sires if sires are partially confounded with groups.

Attention should also be given to the selection of factors in the analyses, since any element of the group has the expected value zero for the corrected sum of squares. Use of an element with expected sum of squares equal to zero would lead to incorrect degrees of freedom in determining mean squares for nonzero elements in an analysis of variance.

Repeatability estimates and most probable producing ability (MPPA) for ratios may be calculated in the normal manner.

SUMMARY

The purposes of this study were to determine whether ratios of a group mean or deviations from a group mean are more appropriate for expressing beef cattle performance records and to extend the method of choice to use in genetic analyses.

The data were 28,545 weaning records collected over a 13-year period (1956-1968) from 203 herds of six breeds in the Iowa Beef Improvement Association program. Only those calves weaned between 160 and 250 days were used in this study. The data were divided into breed, herd, year, season, sex and management groups. There were 2,478 groups with two or more observations.

Statistical properties of ratios were investigated. The assumption that group averages (the denominator of the ratios) were measured without error was investigated. Group variances for ratios were found to be algebraically equal to the square of the coefficients of variation. Reparameterized models using this assumption were found to lead to expected values consistent with the numerical solutions obtained by standard analysis procedures. This assumption was considered to have merit when interpreting ratio data by analysis of variance procedures.

Group averages were proportional to the group variances in sex and management classes and inversely proportional in

season classes. Management and sex were the most important causes of variation in group averages for weaning weight.

Group variances differed more for sex, management, season and their interactions than did the group coefficients of variation for weaning weight. Weaning grade variances were significantly different only for the sex by management effect and coefficients of variation only for management effects.

The expression of data as either ratios or deviations effectively removed group effects due to the group means, so group variances were chosen as the appropriate variable in subsequent analyses. The group averages used in calculating ratios and deviations were regressed to adjust for a small number of observations using the method of Heidhues, *et al*. (1961).

The adjustment of group means for number of observations can change the rank of individuals in different groups and offers advantages in rankings for individuals from groups with high averages.

Subsequent analyses of group variances showed that the variances of ratios differed less than variances of deviations when the variable was weaning weight. The variances of ratios differed significantly only for season effects, where the variances were known to be inversely proportional to the averages. Therefore ratios were considered to be a more appropriate method of expressing weaning weight records than deviations.

Weaning grade variances were least variable for deviated records and no advantage was shown for expressing weaning grades as ratios.

Expected mean squares were developed for different models so that variance components might be estimated. However, genetic variance will likely be underestimated when using either deviations or ratios if sires are partially confounded with groups. If sires are used across groups and have approximately equal numbers of offspring then it may be possible to predict and thus adjust for the loss of genetic variation.

Current estimation procedures for repeatability and most probable producing ability (MPPA) are appropriate when using records expressed as ratios to a group average.

LITERATURE CITED

Bartlett, M. S. 1947. The use of transformations. Biometrics 3: 39.

Brinks, J. S., R. T. Clark, F. J. Rice and N. M. Kieffer. 1961. Adjusting birth weight, weaning weight and preweaning gain for sex of calf in range Hereford cattle. J. Animal Sci. 20: 363.

Cochran, W. G. 1947. Some consequences where the assumptions for the analysis of variance are not satisfied. Biometrics 3: 23.

Cundiff, L. V. 1966. Some factors affecting weaning weights of calves in Oklahoma. Unpublished Ph.D. thesis. Library, Oklahoma State University, Stillwater, Oklahoma.

Eisenhart, Churchill. 1947. The assumptions underlying the analysis of variance. Biometrics 3: 1.

Heidhues, T., L. D. Van Vleck and C. R. Henderson. 1961. Actual and expected accuracy of sire proofs under the New York system of sampling bulls. Zeitschrift fur Tierzuchtung and Zuchtung Biologie 75: 323.

Henderson, C. R. 1948. Estimation of general, specific and maternal combining abilities in crosses among inbred lines of swine. Unpublished Ph.D. thesis. Library, Iowa State University, Ames, Iowa.

Kempthorne, Oscar. 1969. An introduction to genetic statistics. Ames, Iowa, The Iowa State University Press.

Koch, R. M., K. E. Gregory, J. E. Ingalls and R. L. Arthaud. 1959. Evaluating the influence of sex on birth weight and preweaning gain in beef cattle. J. Animal Sci. 18: 738.

Sellers, Halliburton Ian. 1968. The effect of certain factors on weaning weight and weaning grade of beef cattle. Unpublished Ph.D. thesis. Library, Iowa State University, Ames, Iowa.

Steel, Robert G. D. and James H. Torrie. 1960. Principles and procedures of statistics. New York, New York, McGraw-Hill Book Company, Inc.

Van Vleck, L. D., T. Heidhues and C. R. Henderson. 1961. Analysis of deviations of dairy records from different contemporary averages. J. Dairy Sci. 44: 269.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Dr. R. L. Willham for his suggestion of the topic and his continued guidance and counsel during this investigation. The assistance of Dr. E. Pollak for suggestions during the examination of statistical properties is appreciated. The contributions of Dr. Gordon M. Thomson, who assisted with the computations and Dr. R. C. deBaca, who made the data available, are also appreciated.